AD-A263 752

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
3/5/93

3. REPORT TYPE AND DATES COVERED
Final 1 Jul 88 - 30 Jun 91

4. TITLE AND SUBTITLE
Computation and Generalization in Neural Networks

5. FUNDING NUMBERS
DAAL03-88-K-0116

6. AUTHOR(S)
Leon N Cooper

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Brown University
Institute for Brain and Neural Systems
Providence, Rhode Island 02912

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARO 25515.3-LS

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
During this contract period, our research into backward propogation has led to a number of new theoretical and empirical results. We have developed a generalized version of backward propagation. In our generalized network, both gains and synapses are modified by a backward propagation procedure. Synapses are modified in proportion to the negative gradient of the energy with respect to the synaptic weight as in ordinary backward propagation, and gains are modified in proportion to the negative partial derivitive with respect to gain. Since the resulting error signals for the gain and synaptic weights are proportional to one another, the computational complexity of our generalized network is comparable to that of the original backward propagation model.

14. SUBJECT TERMS
Back Propagation
Multilayer Perceptrons
Gain Modification

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

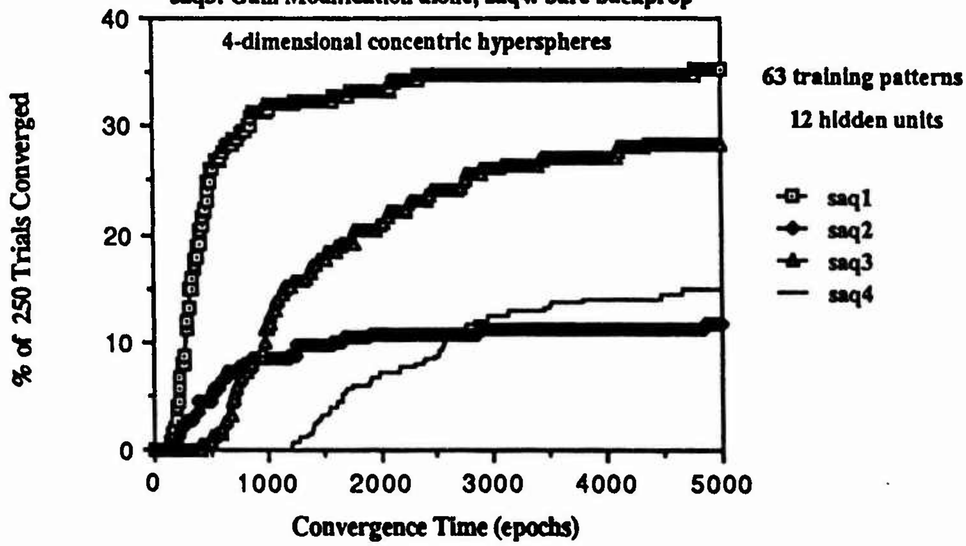
During this contract period, our research into backward propagation has led to a number of new theoretical and empirical results. We have developed a generalized version of backward propagation. In our generalized network, both gains and synapses are modified by a backward propagation procedure. Synapses are modified in proportion to the negative gradient of the energy with respect to the synaptic weight as in ordinary backward propagation, and gains are modified in proportion to the negative partial derivative with respect to the gain. Since the resulting error signals for the gain and synaptic weights are proportional to one another, the computational complexity of our generalized network is comparable to that of the original backward propagation model.

Simulations of the new network have been performed for a concentric circle paradigm in two-dimensions, a concentric sphere paradigm in three dimensions, and a concentric hypersphere paradigm in four dimensions. In the concentric circle problem, we present the x and y coordinates of patterns in the unit circle. Those patterns which lie outside of a predetermined radius are in one class, while those interior to the radius belong to a second class. The concentric sphere and concentric hypersphere paradigms are analogous to the concentric circle problem, except that they are in three and higher dimensions respectively.

In simulations, the convergence rate of backward propagation is found to increase when the momentum is increased. When we use gain modification in addition to high momentum, the convergence rate is still faster. For the concentric circle problem, we have found that the number of epochs required for 80% of the trials to converge to a solution is 6-7 times smaller when high momentum is used. The 80% convergence level for trials using both gain modification and high-momentum synaptic modification is reached in about 1/3 of the time required when using high-momentum synaptic modification alone. These improvements in convergence rate are due to the more rapid development of the effective synaptic vectors in the network ; the effective synaptic vectors are defined as the product of the ordinary synaptic vector and the gain (in our generalized network, the gains are initialized to unity, but evolve in time; in ordinary backward propagation the gains are all set equal to unity). Theoretically, we have shown that gain modification is equivalent to the use of an effective time-dependent step-constant for the rescaled synapses. In the synaptic vector space, the step-constant has a quadratic dependence on the magnitude of the ordinary synaptic vector and also has a quadratic dependence on the gain of the cell in the direction of the rescaled synapses. In the hyperplane perpendicular to the rescaled synapses, the step constant depends quadratically on the gain alone. The theoretical development of gain modification and the effective time dependent step-constant, as well as the results of the simulations of the two-dimensional concentric circle problem, are detailed in our recent technical report (Bachmann, ARO technical report #11, December, 1989).

Experiments to chart the improvements due to gain modification as a function of dimensionality are under way. In our current studies, we use higher dimensional concentric hyperspheres. Results have been obtained in three and four dimensions which further demonstrate the improved convergence rate obtained by using a combination of gain modification and high momentum. An example of convergence curves for a four-dimensional concentric hypersphere paradigm appear in the graph below

Saq1: Gain Modification + High Momentum; saq2: High Momentum alone;
saq3: Gain Modification alone; saq4: bare backprop



Our recent work has also focussed on the development of an algorithm for scheduling the presentation of patterns in backward propagation. Ahmad and Tesauro (1988) have suggested that the use of class boundary patterns in training neural networks can greatly enhance the convergence rate and generality of learning in backward propagation networks. Their work concerned the simple majority-function paradigm for binary inputs. Patterns one hamming unit away from the class boundary were used for training. An algorithm which we have formulated preprocesses patterns to determine how close the training patterns are to a class boundary. For each pattern, we compute the minimum distance to a pattern of the opposite class. In one version of our algorithm, we then divide the training patterns into two zones on the basis of this distance metric. Those patterns falling in the "near zone", i.e. those closest to patterns of the opposite class, are eligible for modification every time, provided they are above the modification threshold. Those in the "far zone" are eligible at a lower freequency. One can generalize this approach to multiple zones, of varying width and eligibility frequency. An alternative approach is to consider mapping a continuous monotonically increasing eligiblity function on the range defined by this distance metric. The eligibility time-step for any particular pattern is then defined by computing the non-linear function and rounding to the nearest integer. This effectively amounts to creating multiple zones, but choosing the widths and associated frequencies according to a particular rule. Details have been presented in technical reports and publications.